

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Bandman, Olga
Corley, Neil C.
Guegler, Karl J.

(ii) TITLE OF THE INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
(B) STREET: 3174 Porter Drive
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: To Be Assigned
(B) FILING DATE: Herewith
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Billings, Lucy J.
(B) REGISTRATION NUMBER: 36,749
(C) REFERENCE/DOCKET NUMBER: PF-0333 US

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 415-855-0555
(B) TELEFAX: 415-845-4166
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 448 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: CORNNOT01
(B) CLONE: 45517

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met Pro Gly Ile Lys Arg Ile Leu Thr Val Thr Ile Leu Ala Leu Cys
 1 5 10 15
 Leu Pro Ser Pro Gly Asn Ala Gln Ala Gln Cys Thr Asn Gly Phe Asp
 20 25 30
 Leu Asp Arg Gln Ser Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Thr
 35 40 45
 Ile Pro Glu Ala Cys Arg Gly Asp Met Met Cys Val Asn Gln Asn Gly
 50 55 60
 Gly Tyr Leu Cys Ile Pro Arg Thr Asn Pro Val Tyr Arg Gly Pro Tyr
 65 70 75 80
 Ser Asn Pro Tyr Ser Thr Pro Tyr Ser Gly Pro Tyr Pro Ala Ala Ala
 85 90 95
 Pro Pro Leu Ser Ala Pro Asn Tyr Pro Thr Ile Ser Arg Pro Leu Ile
 100 105 110
 Cys Arg Phe Gly Tyr Gln Met Asp Glu Ser Asn Gln Cys Val Asp Val
 115 120 125
 Asp Glu Cys Ala Thr Asp Ser His Gln Cys Asn Pro Thr Gln Ile Cys
 130 135 140
 Ile Asn Thr Glu Gly Tyr Thr Cys Ser Cys Thr Asp Gly Tyr Trp
 145 150 155 160
 Leu Leu Glu Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Tyr Gly Tyr
 165 170 175
 Cys Gln Gln Leu Cys Ala Asn Val Pro Gly Ser Tyr Ser Cys Thr Cys
 180 185 190
 Asn Pro Gly Phe Thr Leu Asn Glu Asp Gly Arg Ser Cys Gln Asp Val
 195 200 205
 Asn Glu Cys Ala Thr Glu Asn Pro Cys Val Gln Thr Cys Val Asn Thr
 210 215 220
 Tyr Gly Ser Phe Ile Cys Arg Cys Asp Pro Gly Tyr Glu Leu Glu Glu
 225 230 235 240
 Asp Gly Val His Cys Ser Asp Met Asp Glu Cys Ser Phe Ser Glu Phe
 245 250 255
 Leu Cys Gln His Glu Cys Val Asn Gln Pro Gly Thr Tyr Phe Cys Ser
 260 265 270
 Cys Pro Pro Gly Tyr Ile Leu Leu Asp Asp Asn Arg Ser Cys Gln Asp
 275 280 285
 Ile Asn Glu Cys Glu His Arg Asn His Thr Cys Asn Leu Gln Gln Thr
 290 295 300
 Cys Tyr Asn Leu Gln Gly Gly Phe Lys Cys Ile Asp Pro Ile Arg Cys
 305 310 315 320
 Glu Glu Pro Tyr Leu Arg Ile Ser Asp Asn Arg Cys Met Cys Pro Ala
 325 330 335
 Glu Asn Pro Gly Cys Arg Asp Gln Pro Phe Thr Ile Leu Tyr Arg Asp
 340 345 350
 Met Asp Val Val Ser Gly Arg Ser Val Pro Ala Asp Ile Phe Gln Met
 355 360 365
 Gln Ala Thr Thr Arg Tyr Pro Gly Ala Tyr Tyr Ile Phe Gln Ile Lys
 370 375 380
 Ser Gly Asn Glu Gly Arg Glu Phe Tyr Met Arg Gln Thr Gly Pro Ile
 385 390 395 400
 Ser Ala Thr Leu Val Met Thr Arg Pro Ile Lys Gly Pro Arg Glu Ile
 405 410 415
 Gln Leu Asp Leu Glu Met Ile Thr Val Asn Thr Val Ile Asn Phe Arg
 420 425 430
 Gly Ser Ser Val Ile Arg Leu Arg Ile Tyr Val Ser Gln Tyr Pro Phe
 435 440 445

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2550 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (vii) IMMEDIATE SOURCE:
- (A) LIBRARY: CORNNOT01
 - (B) CLONE: 45517

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CCAAGATTGT	TGTGAGGAGT	CTAGCCAGTT	GGTGAGCGCT	GTAATCTGAA	CCAGCTGTGT	60
CCAGACTGAG	GCCCCATTG	CATTATTAA	CATACTTAGA	AAATGAAGTG	TTCATTTTA	120
ACATTCTCC	TCCAATTGGT	TTAATGCTGA	ATTACTGAAG	AGGGCTAACG	AAAACCAGGT	180
GCTTGCCTG	AGGGCTCTGC	AGTGGCTGGG	AGGACCCCGG	CGCTCTCCCC	GTGTCCTCTC	240
CACGACTCGC	TGGGCCCTC	TGGAATAAAA	CACCCGCGAG	CCCCGAGGGC	CCAGAGGAGG	300
CCGACGTGCC	CGAGCTCCTC	CGGGGGTCCC	CCCCGCGAGC	TTTCTTCTCG	CCTTCGCATC	360
TCCTCTCGC	CGCTCTTGGA	CATGCCAGGA	ATAAAAAAGGA	TACTCACTGT	TACCATTCTG	420
GCTCTCTGTC	TTCCAAGCCC	TGGGAATGCA	CAGGCACAGT	GCACGAATGG	CTTGACCTG	480
GATGCCAGT	CAGGACAGTG	TTTAGATATT	GATGAATGCC	GAACCATCCC	CGAGGCCTGC	540
CGAGGAGACA	TGATGTGTGT	TAACCAAAAT	GGCGGGTATT	TATGCATTCC	CCGGACAAAC	600
CCTGTGTATC	GAGGGCCCTA	CTCGAACCCC	TACTCGACCC	CCTACTCAGG	TCCGTACCCA	660
GCAGCTGCC	CACCACTCTC	AGCTCCAAAC	TATCCCACGA	TCTCCAGGCC	TCTTATATGC	720
CGCTTGGAT	ACCAGATGGA	TGAAAGCAAC	CAATGTGTGG	ATGTTGGACGA	GTGTGCAACA	780
GATTCCCACC	AGTGAACCCC	CACCCAGATC	TGCATCAATA	CTGAAGGCGG	GTACACCTGC	840
TCCTGCACCG	ACGGATATTG	GCTCTGGAA	GGCCAGTGT	TAGACATTGA	TGAATGTCCG	900
TATGGTTACT	CCCAGCAGCT	CTGTGCGAAT	GTTCCTGGAT	CCTATTCTTG	TACATGCAAC	960
CCTGGTTTA	CCCTCAATGA	GGATGGAAGG	TCTTGCCAAG	ATGTGAACGA	GTGTGCCACC	1020
GAGAACCCCT	CGCTGCAAAC	CTGCGTCAAC	ACCTACGGCT	CTTTCATCTG	CCGCTGTGAC	1080
CCAGGATATG	AACTTGAGGA	AGATGGCGTT	CATTGCAGTG	ATATGGACGA	GTGCAGCTTC	1140
TCTGAGTTCC	TCTGCCAAC	TGAGTGTGTG	AACCAGCCCG	GCACATACTT	CTGCTCTGC	1200
CCTCCAGGCT	ACATCCTGCT	GGATGACAAC	CGAAGCTGCC	AAGACATCAA	CGAATGTGAG	1260
CACAGGAACC	ACACGTGCAA	CCTGCAGCAG	ACGTGCTACA	ATTTACAAGG	GGGCTTCAAA	1320
TGCATCGACC	CCATCCGCTG	TGAGGAGCCT	TATCTGAGGA	TCAGTGATAA	CCGCTGTATG	1380
TGTCCCTGCTG	AGAACCCCTGG	CTGCAGAGAC	CAGCCCTTA	CCATCTTGT	CCGGGACATG	1440
GACGTTGGTGT	CAGGACGCTC	CGTTCCCGCT	GACATCTTCC	AAATGCAAGC	CACGACCCGC	1500
TACCTGGGG	CCTATTACAT	TTTCCAGATC	AAATCTGGGA	ATGAGGGCAG	AGAATTTTAC	1560
ATGCGGCCAA	CGGGCCCCAT	CAGTGCCACC	CTGGTGATGA	CACGCCCAT	CAAAGGGCCC	1620
CGGGAATCC	AGCTGGACTT	GGAAATGATC	ACTGTCAACA	CTGTCTCAA	CTTCAGAGGC	1680
AGCTCCGTGA	TCCGACTGCG	GATATATGTG	TCGCACTTAC	CATTCTGAGC	CTCGGGCTGG	1740
AGCCTCCGAC	GCTGCCCTCTC	ATTGGCACCA	AGGGACAGGA	GAAGAGAGGA	AATAACAGAG	1800
AGAATGAGAG	CGACACAGAC	GTTAGGCATT	TCCCTGCTGA	CGTTCCCG	AAGAGTCAGC	1860
CCCGACTTCC	TGACTCTCAC	CTGTACTATT	GCAGACCTGT	CACCTGCAG	GACTTGCAC	1920
CCCCAGTTCC	TATGACACAG	TTATCAAAAA	GTATTATCAT	TGCTCCCTG	ATAGAAGATT	1980
GTTGGTGAAT	TTTCAAGGCC	TTCAGTTAT	TTCCACTATT	TTCAAAGAAA	ATAGATTAGG	2040
TTTGGGGGG	TCTGAGTCTA	TGTTCAAAGA	CTGTGAACAG	CTTGTGTCA	CTTCTTCACC	2100
TCTTCCACTC	CTTCTCTCAC	TGTGTTACTG	CTTTGCAAAG	ACCCGGGAGC	TGGCGGGGAA	2160
CCCTGGGAGT	AGCTAGTTG	CTTTTGCCT	ACACAGAGAA	GGCTATGTAA	ACAAACACAA	2220
GCAGGATCGA	AGGGTTTTA	GAGAATGTGT	TTCAAAACCA	TGCCTGGTAT	TTTCAACCCAT	2280
AAAAGAAGTT	TCAGTTGTCC	TTAAATTTGT	ATAACGGTT	AATTCTGTCT	TGTTCATTTT	2340
GAGTATTTTT	AAAAAATATG	TCGTAGAATT	CCTTCGAAAG	GCCTTCAGAC	ACATGCTATG	2400
TTCTGTCTTC	CCAAACCCAG	TCTCCTCTCC	ATTAGGCCTC	AGTGTGTTCT	TTGAGGACCC	2460
CTTAATCTTG	CTTTCTTTAG	AATTGTTACC	CAATTGGATT	GGAATGCGAGA	GGTCTCCAAA	2520
CTGATTAAAT	ATTGAGAG	AAAAAA				2550

- (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 540 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
 (A) LIBRARY: BRAITUT13
 (B) CLONE: 1621777

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Gly Thr Thr Ala Arg Ala Ala Leu Val Leu Thr Tyr Leu Ala Val
 1 5 10 15
 Ala Ser Ala Ala Ser Glu Gly Gly Phe Thr Ala Thr Gly Gln Arg Gln
 20 25 30
 Leu Arg Pro Glu His Phe Gln Glu Val Gly Tyr Ala Ala Pro Pro Ser
 35 40 45
 Pro Pro Leu Ser Arg Ser Leu Pro Met Asp His Pro Asp Ser Ser Gln
 50 55 60
 His Gly Pro Pro Phe Glu Gly Gln Ser Gln Val Gln Pro Pro Pro Ser
 65 70 75 80
 Gln Glu Ala Thr Pro Leu Gln Gln Glu Lys Leu Leu Pro Ala Gln Leu
 85 90 95
 Pro Ala Glu Lys Glu Val Gly Pro Pro Leu Pro Gln Glu Ala Val Pro
 100 105 110
 Leu Gln Lys Glu Leu Pro Ser Leu Gln His Pro Asn Glu Gln Lys Glu
 115 120 125
 Gly Met Pro Ala Pro Phe Gly Asp Gln Ser His Pro Glu Pro Glu Ser
 130 135 140
 Trp Asn Ala Ala Gln His Cys Gln Gln Asp Arg Ser Gln Gly Gly Trp
 145 150 155 160
 Gly His Arg Leu Asp Gly Phe Pro Pro Gly Arg Pro Ser Pro Asp Asn
 165 170 175
 Leu Asn Gln Ile Cys Leu Pro Asn Arg Gln His Val Val Tyr Gly Pro
 180 185 190
 Trp Asn Leu Pro Gln Ser Ser Tyr Ser His Leu Thr Arg Gln Gly Glu
 195 200 205
 Thr Leu Asn Phe Leu Glu Ile Gly Tyr Ser Arg Cys Cys His Cys Arg
 210 215 220
 Ser His Thr Asn Arg Leu Glu Cys Ala Lys Leu Val Trp Glu Glu Ala
 225 230 235 240
 Met Ser Arg Phe Cys Glu Ala Glu Phe Ser Val Lys Thr Arg Pro His
 245 250 255
 Trp Cys Cys Thr Arg Gln Gly Glu Ala Arg Phe Ser Cys Phe Gln Glu
 260 265 270
 Glu Ala Pro Gln Pro His Tyr Gln Leu Arg Ala Cys Pro Ser His Gln
 275 280 285
 Pro Asp Ile Ser Ser Gly Leu Glu Leu Pro Phe Pro Pro Gly Val Pro
 290 295 300
 Thr Leu Asp Asn Ile Lys Asn Ile Cys His Leu Arg Arg Phe Arg Ser
 305 310 315 320
 Val Pro Arg Asn Leu Pro Ala Thr Asp Pro Leu Gln Arg Glu Leu Leu
 325 330 335
 Ala Leu Ile Gln Leu Glu Arg Glu Phe Gln Arg Cys Cys Arg Gln Gly
 340 345 350
 Asn Asn His Thr Cys Thr Trp Lys Ala Trp Glu Asp Thr Leu Asp Lys
 355 360 365
 Tyr Cys Asp Arg Glu Tyr Ala Val Lys Thr His His His Leu Cys Cys
 370 375 380
 Arg His Pro Pro Ser Pro Thr Arg Asp Glu Cys Phe Ala Arg Arg Ala
 385 390 395 400
 Pro Tyr Pro Asn Tyr Asp Arg Asp Ile Leu Thr Ile Asp Ile Gly Arg
 405 410 415
 Val Thr Pro Asn Leu Met Gly His Leu Cys Gly Asn Gln Arg Val Leu
 420 425 430

Thr	Lys	His	Lys	His	Ile	Pro	Gly	Leu	Ile	His	Asn	Met	Thr	Ala	Arg
435					440					445					
Cys	Cys	Asp	Leu	Pro	Phe	Pro	Glu	Gln	Ala	Cys	Cys	Ala	Glu	Glu	Glu
450					455					460					
Lys	Leu	Thr	Phe	Ile	Asn	Asp	Leu	Cys	Gly	Pro	Arg	Arg	Asn	Ile	Trp
465					470					475			480		
Arg	Asp	Pro	Ala	Leu	Cys	Cys	Tyr	Leu	Ser	Pro	Gly	Asp	Glu	Gln	Val
					485				490			495			
Asn	Cys	Phe	Asn	Ile	Asn	Tyr	Leu	Arg	Asn	Val	Ala	Leu	Val	Ser	Gly
					500				505			510			
Asp	Thr	Glu	Asn	Ala	Lys	Gly	Gln	Gly	Glu	Gln	Gly	Ser	Thr	Gly	Gly
					515				520			525			
Thr	Asn	Ile	Ser	Ser	Thr	Ser	Glu	Pro	Lys	Glu	Glu				
					530				535			540			

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1899 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: BRAITUT13
- (B) CLONE: 162177

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TGGGTGCAAG	CTCACACCCG	TAACAGCCAC	CAGACAAGCT	TCAGTGGCCG	GCCCTTCACA	60
TCCAGACTTG	CCTGAGAGGA	CCCACCTCTG	AGTGTCCAGT	GGTCAGTTGC	CCCAGGATGG	120
GGACCACAGC	CAGAGCAGCC	TTGGTCTTGA	CCTATTGCG	TGTTGCTTCT	GCTGCCCTTG	180
AGGGAGGCTT	CACGGCTACA	GGACAGAGGC	AGCTGAGGCC	AGAGCACTTT	CAAGAAGTTG	240
GCTACCGAGC	TCCCCCCTCC	CCACCCCTAT	CCCGAAGCCT	CCCCATGGAT	CACCCCTGACT	300
CCTCTCAGCA	TGGCCCTCCC	TTTGAGGGAC	AGAGTCAGT	GCAGCCCCCT	CCCTCTCAGG	360
AGGCCACCCC	TCTCCAACAG	GAAAAGCTGC	TACCTGCCCA	ACTCCCTGCT	GAAAAGGAAG	420
TGGGTCCCCC	TCTCCCTCAG	GAAGCTGTCC	CCCTCCAAA	AGAGCTGCC	TCTCTCCAGC	480
ACCCCAATGA	ACAGAAGGAA	GBAATGCCAG	CTCCATTGG	GGACCAAGAGC	CATCCAGAAC	540
CTGAGTCCTG	GAATGCAGCC	CAGCACTGCC	AACAGGACCG	GTCCAAGGG	GGCTGGGCC	600
ACCGGCTGGA	TGGCTCCCC	CCTGGCGGC	CTTCTCCAGA	CAATCTGAAC	CAAATCTGCC	660
TTCCTAACCG	TCAGCATGTG	GTATATGGTC	CCTGGAACCT	ACCACAGTCC	AGCTACTCCC	720
ACCTCACTCG	CCAGGGTGAG	ACCCCTCAATT	TCCTGGAGAT	TGGTATTCC	CGCTGCTGCC	780
ACTGCCGAG	CCACACAAAC	CGCCTAGAGT	GTGCCAAACT	TGTGTTGGAG	GAAGCAATGA	840
GCCGATTCTG	TGAGGCCGAG	TTCTCGGTCA	AGACCCGACC	CCACTGGTGC	TGCACGGGC	900
AGGGGGAGGC	TCGGTTCTCC	TGCTTCCAGG	AGGAAGCTCC	CCAGCCACAC	TACCAGCTCC	960
GGGCCTGCC	CAGCCATCAG	CCTGATATT	CCTCGGGTCT	TGAGCTGCC	TTCCCTCCTG	1020
GGGTGCCAC	ATTGGACAAT	ATCAAGAAC	TCTGCCACT	GAGGCCTTC	CGCTCTGTG	1080
CACGCCACCT	GCCAGCTACT	GACCCCTAC	AAAGGGAGCT	GCTGGCACTG	ATCCAGCTGG	1140
AGAGGGAGTT	CCAGCGCTGC	TGCCGCCAGG	GGAACAATCA	CACCTGTACA	TGGAAGGCCT	1200
GGGAGGATAC	CCTTGACAAA	TACTGTGACC	GGGAGTATGC	TGTGAAGACC	CACCACCACT	1260
TGTGTTGCCG	CCACCCCTCC	AGCCCTACTC	GGGATGAGTG	CTTTGCCCGT	CGGGCTCCTT	1320
ACCCCAACTA	TGACCGGGAC	ATCTTGACCA	TTGACATCGG	TCGAGTCACC	CCCAACCTCA	1380
TGGGCCACCT	CTGTGGAAAC	CAAAGAGTTC	TCACCAAGCA	TAAACATATT	CCTGGGCTGA	1440
TCCACAAACAT	GACTGCCCGC	TGCTGTGACC	TGCCATTTC	AGAACAGGCC	TGCTGTGCAG	1500
AGGAGGAGAA	ATTAACCTTC	ATCAATGATC	TGTGTGGTCC	CCGACGTAAC	ATCTGGCGAG	1560
ACCCCTGCCCT	CTGCTGTTAC	CTGAGTCCTG	GGGATGAACA	GGTCAACTGC	TTCAACATCA	1620
ATTATCTGAG	GAACGTGGCT	CTAGTGTCTG	GAGACACTGA	GAACGCCAAG	GGCCAGGGGG	1680
AGCAGGGCTC	AACTGGAGGA	ACAAATATCA	GCTCCACCTC	TGAGCCCAAG	GAAGAATGAG	1740
TCACCCAGA	GCCCTAGAGG	GTCAGATGGG	GGGAACCCCA	CCCTGCCCTCA	CCCATCTGAA	1800
CACTCATTAC	ACTAAACACC	TCTTGGATT	GGTGTCTCA	TTGTCTATCT	AATGTCTCAC	1860
CCGCAGTGT	TTAAGTGGAT	CTTGGTGC	GGCCAGG			1899

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 387 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 458228

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Ala Thr Ser Gly Val Leu Pro Gly Gly Gly Phe Val Ala Ser Ala
 1 5 10 15
 Ala Ala Val Ala Gly Pro Glu Met Gln Thr Gly Arg Asn Asn Phe Val
 20 25 30
 Ile Arg Arg Asn Pro Ala Asp Pro Gln Arg Ile Pro Ser Asn Pro Ser
 35 40 45
 His Arg Ile Gln Cys Ala Ala Gly Tyr Glu Gln Ser Glu His Asn Val
 50 55 60
 Cys Gln Asp Ile Asp Glu Cys Thr Ala Gly Thr His Asn Cys Arg Ala
 65 70 75 80
 Asp Gln Val Cys Ile Asn Leu Arg Gly Ser Phe Ala Cys Gln Cys Pro
 85 90 95
 Pro Gly Tyr Gln Lys Arg Gly Glu Gln Cys Val Asp Ile Asp Glu Cys
 100 105 110
 Thr Ile Pro Pro Tyr Cys His Gln Arg Cys Val Asn Thr Pro Gly Ser
 115 120 125
 Phe Tyr Cys Gln Cys Ser Pro Gly Phe Gln Leu Ala Ala Asn Asn Tyr
 130 135 140
 Thr Cys Val Asp Ile Asn Glu Cys Asp Ala Ser Asn Gln Cys Ala Gln
 145 150 155 160
 Gln Cys Tyr Asn Ile Leu Gly Ser Phe Ile Cys Gln Cys Asn Gln Gly
 165 170 175
 Tyr Glu Leu Ser Ser Asp Arg Leu Asn Cys Glu Asp Ile Asp Glu Cys
 180 185 190
 Arg Thr Ser Ser Tyr Leu Cys Gln Tyr Gln Cys Val Asn Glu Pro Gly
 195 200 205
 Lys Phe Ser Cys Met Cys Pro Gln Gly Tyr Gln Val Val Arg Ser Arg
 210 215 220
 Thr Cys Gln Asp Ile Asn Glu Cys Glu Thr Thr Asn Glu Cys Arg Glu
 225 230 235 240
 Asp Glu Met Cys Trp Asn Tyr His Gly Gly Phe Arg Cys Tyr Pro Arg
 245 250 255
 Asn Pro Cys Gln Asp Pro Tyr Ile Leu Thr Pro Glu Asn Arg Cys Val
 260 265 270
 Cys Pro Val Ser Asn Ala Met Cys Arg Glu Leu Pro Gln Ser Ile Val
 275 280 285
 Tyr Lys Tyr Met Ser Ile Arg Ser Asp Arg Ser Val Pro Ser Asp Ile
 290 295 300
 Phe Gln Ile Gln Ala Thr Thr Ile Tyr Ala Asn Thr Ile Asn Thr Phe
 305 310 315 320
 Arg Ile Lys Ser Gly Asn Glu Asn Gly Glu Phe Tyr Leu Arg Gln Thr
 325 330 335
 Ser Pro Val Ser Ala Met Leu Val Leu Val Lys Ser Leu Ser Gly Pro
 340 345 350
 Arg Glu His Ile Val Asp Leu Glu Met Leu Thr Val Ser Ser Ile Gly
 355 360 365

Thr Phe Arg Thr Ser Ser Val Leu Arg Leu Thr Ile Ile Val Gly Pro
 370 375 380
 Phe Ser Phe
 385

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 559 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 496120

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Gly Thr Val Ser Arg Ala Ala Leu Ile Leu Ala Cys Leu Ala Leu
 1 5 10 15
 Ala Ser Ala Ala Ser Glu Gly Ala Phe Lys Ala Ser Asp Gln Arg Glu
 20 25 30
 Met Thr Pro Glu Arg Leu Phe Gln His Leu His Glu Val Gly Tyr Ala
 35 40 45
 Ala Pro Pro Ser Leu Pro Gln Thr Arg Arg Leu Arg Val Asp His Ser
 50 55 60
 Val Thr Ser Leu His Asp Pro Pro Leu Phe Glu Glu Gln Arg Glu Val
 65 70 75 80
 Gln Pro Pro Ser Ser Pro Glu Asp Ile Pro Val Tyr Glu Glu Asp Trp
 85 90 95
 Pro Thr Phe Leu Asn Pro Asn Val Asp Lys Ala Gly Pro Ala Val Pro
 100 105 110
 Gln Glu Ala Ile Pro Leu Gln Lys Glu Gln Pro Pro Gln Val His
 115 120 125
 Ile Glu Gln Lys Glu Ile Asp Pro Pro Ala Gln Pro Gln Glu Glu Ile
 130 135 140
 Val Gln Lys Glu Val Lys Pro His Thr Leu Ala Gly Gln Leu Pro Pro
 145 150 155 160
 Glu Pro Arg Thr Trp Asn Pro Ala Arg His Cys Gln Gln Gly Arg Arg
 165 170 175
 Gly Val Trp Gly His Arg Leu Asp Gly Phe Pro Pro Gly Arg Pro Ser
 180 185 190
 Pro Asp Asn Leu Lys Gln Ile Cys Leu Pro Glu Arg Gln His Val Ile
 195 200 205
 Tyr Gly Pro Trp Asn Leu Pro Gln Thr Gly Tyr Ser His Leu Ser Arg
 210 215 220
 Gln Gly Glu Thr Leu Asn Val Leu Glu Thr Gly Tyr Ser Arg Cys Cys
 225 230 235 240
 Pro Cys Arg Ser Asp Thr Asn Arg Leu Asp Cys Leu Lys Leu Val Trp
 245 250 255
 Glu Asp Ala Met Thr Gln Phe Cys Glu Ala Glu Phe Ser Val Lys Thr
 260 265 270
 Arg Pro His Leu Cys Cys Arg Leu Arg Gly Glu Glu Arg Phe Ser Cys
 275 280 285
 Phe Gln Lys Glu Ala Pro Arg Pro Asp Tyr Leu Leu Arg Pro Cys Pro
 290 295 300
 Val His Gln Asn Gly Met Ser Ser Gly Pro Gln Leu Pro Phe Pro Pro
 305 310 315 320
 Gly Leu Pro Thr Pro Asp Asn Val Lys Asn Ile Cys Leu Leu Arg Arg
 325 330 335

Phe Arg Ala Val Pro Arg Asn Leu Pro Ala Thr Asp Ala Ile Gln Arg
 340 345 350
 Gln Leu Gln Ala Leu Thr Arg Leu Glu Thr Glu Phe Gln Arg Cys Cys
 355 360 365
 Arg Gln Gly His Asn His Thr Cys Thr Trp Lys Ala Trp Glu Gly Thr
 370 375 380
 Leu Asp Gly Tyr Cys Glu Arg Glu Leu Ala Ile Lys Thr His Pro His
 385 390 395 400
 Ser Cys Cys His Tyr Pro Pro Ser Pro Ala Arg Asp Glu Cys Phe Ala
 405 410 415
 His Leu Ala Pro Tyr Pro Asn Tyr Asp Arg Asp Ile Leu Thr Leu Asp
 420 425 430
 Leu Ser Arg Val Thr Pro Asn Leu Met Gly Gln Leu Cys Gly Ser Gly
 435 440 445
 Arg Val Leu Ser Lys His Lys Gln Ile Pro Gly Leu Ile Gln Asn Met
 450 455 460
 Thr Val Arg Cys Cys Glu Leu Pro Tyr Pro Glu Gln Ala Cys Cys Gly
 465 470 475 480
 Glu Glu Glu Lys Leu Ala Phe Ile Glu Asn Leu Cys Gly Pro Arg Arg
 485 490 495
 Asn Ser Trp Lys Asp Pro Ala Leu Cys Cys Asp Leu Ser Pro Glu Asp
 500 505 510
 Lys Gln Ile Asn Cys Phe Asn Thr Asn Tyr Leu Arg Asn Val Ala Leu
 515 520 525
 Val Ala Gly Asp Thr Gly Asn Ala Thr Gly Leu Gly Glu Gln Gly Pro
 530 535 540
 Thr Arg Gly Thr Asp Ala Asn Pro Ala Pro Gly Ser Lys Glu Glu
 545 550 555